REMARKS

The invention relates to a transgenic mouse comprising an isolated DNAc molecule comprising a promoter P and an L1 cassette sequence comprising a core L1 retrotransposon element and methods of use thereof.

Claims 34, 36-44, 46, 47 and 49 are currently pending. Claims 34, and 36-43 have been amended to recite a high frequency, insertional mutagenesis model. Support for this amendment can at least be found beginning on page 27, line 7. No new matter has been added by way of this amendment.

Rejection of claims 34, 36-44, 46, 47 and 49 under 35 U.S.C. § 112, first paragraph

The Examiner has maintained the rejection of claims 34, 36-44, 46, 47 and 49 pursuant to 35 U.S.C. §112, first paragraph, for failing to comply with the enablement requirement. Specifically, the Examiner is of the opinion that the specification fails to provide an enabling disclosure for the claimed transgenic mouse because the phenotype of the mouse is unpredictable, and therefore, in the Examiner's view, the specification does not teach how to use the claimed transgenic mouse.

Applicants disagree. Contrary to the Examiner's assertion, the high frequency, insertional mutagenesis model, wherein the model is a transgenic mouse of the invention is enabled by the specification on page 27, lines 22-29 to page 28, line 1. Contrary to the Examiner's assertion in the Office Action at page 4 that the mouse cannot be used to generate mutations in a cell, this section of the specification discloses that this is precisely the use for the mouse. The transgenic mice are generated so that sperm can be obtained therefrom, wherein the sperm are then used to generate high frequency mutations. The transgenic mouse is enabled, has utility and nothing more is required. However, to bolster Applicants' argument, an article published by Dr Kazazian and others after the filing date of the present application is enclosed (Ostertag et al., 2002, Nature Genetics 32:655-660. This article discloses a transgenic mouse made according to the methods described in the present specification having the use asserted in the specification. Applicants point out that this reference has already been submitted to the Examiner, but is enclosed herewith for the Examiner's convenience.

The high frequency, insertional mutagenesis model allows for assessing the mutagenic potential of an animal by assessing the frequency of retrotransposition in the cells of

that animal. Since L1 elements are capable of random insertion into a cell genome, they are mutagenic. Thus, retrotransposition may be used as a measure of mutagenic potential in an animal.

Applicants respectfully submit that the claims, as amended and presently under consideration, are enabled and Applicants respectfully request reconsideration and withdrawal of the Examiner's rejection pursuant to 35 U.S.C. §112, first paragraph, for lack of enablement.

Summary

Applicants respectfully submit that each rejection of the Examiner to the claims of the present application has been overcome or is now inapplicable, and that claims 34 and 36-44, 46, 47, and 49 are now in condition for allowance. Applicants further submit that no new matter has been added by way of the present amendment. Reconsideration and allowance of these claims is respectfully requested at the earliest possible date.

Respectfully submitted,

Haig Kazazian *et al.*

October 31, 2007
(Date)

By: _____

QUANG D. NGUYEN, Ph.D.
Registration No. 52,066
DRINKER, BIDDLE & REATH, LLP
One Logan Square
18th and Cherry Streets
Philadelphia, PA 19103-6996
Telephone: (215) 988-2700
Direct Dial: (215) 988-2720
Facsimile: (215) 988-2757
E-Mail: Quang.Nguyen@dbr.com
Agent for Applicants

Enclosures: Request for Continued Examination
Ostertag et al., 2002, Nature Genetics 32:655-660

dcl

A mouse model of human L1 retrotransposition

Eric M. Ostertag[1], Ralph J. DeBerardinis[1], John L. Goodier[1], Yue Zhang[1], Nuo Yang[1], George L. Gerton[2] & Haig H. Kazazian Jr.[1]

Published online 4 November 2002; doi:10.1038/ng1022

The L1 retrotransposon has had an immense impact on the size and structure of the human genome through a variety of mechanisms, including insertional mutagenesis[1,2]. To study retrotransposition in a living organism, we created a mouse model of human L1 retrotransposition. Here we show that L1 elements can retrotranspose in male germ cells, and that expression of a human L1 element under the control of its endogenous promoter is restricted to testis and ovary. In the mouse line with the highest level of L1 expression, we found two *de novo* L1 insertions in 135 offspring. Both insertions were structurally indistinguishable from natural endogenous insertions. This suggests that an individual L1 element can have substantial mutagenic potential. In addition to providing a valuable *in vivo* model of retrotransposition in mammals, these mice are an important step in the development of a new random mutagenesis system.

L1 retrotransposons are directly or indirectly responsible for an estimated one-third of the human genome[1]. Of the approximately 500,000 L1 sequences in the human diploid genome, 40–80 remain active (that is, retrotranspositionally competent) and are a source of *de novo* mutations in humans[1–3]. Although a number of these recent L1 insertions have been characterized, it has not yet been possible to determine in which cells or when these events occur.

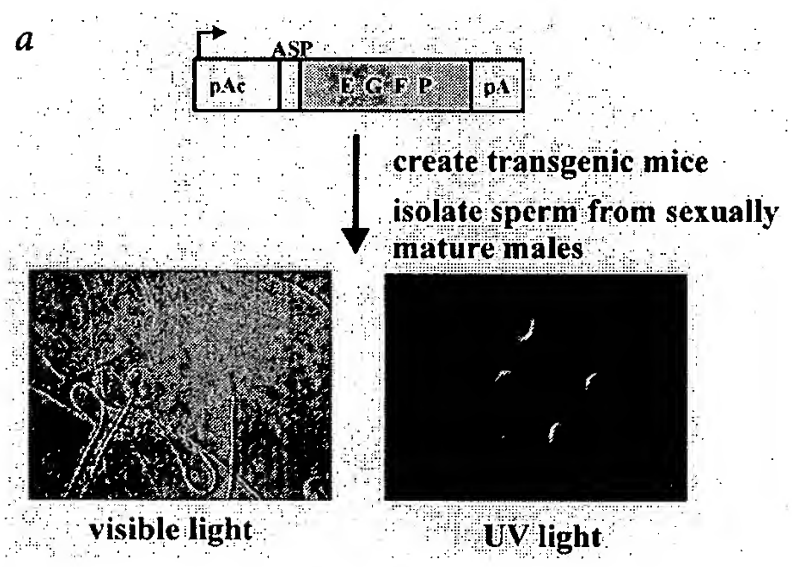

Fig. 1 Transgenic mouse lines. *a*, The transgene used to create the GFP-positive control line consisted of a modified EGFP gene that contained a preproacrosin promoter (pAc and bent arrow) and an acrosin signal peptide (ASP). Acrosomes of sperm from sexually mature males of this line fluoresced under ultraviolet (UV) light (right panel). The left panel shows the same sperm under visible light. *b*, The modified EGFP-based retrotransposition cassette. The L1 element consists of 5′ and 3′ UTRs and two open reading frames (ORFs). The 5′ UTR contains an internal promoter[14], ORF1 encodes an RNA binding protein[27,28], and ORF2 encodes an endonuclease (EN)[20], a reverse transcriptase (RT)[29] and a conserved cysteine-rich domain (C)[30]. The EGFP retrotransposition cassette was cloned into the L1 3′ UTR in the antisense orientation. It consisted of the preproacrosin promoter (pAc) and acrosin signal peptide (ASP), the EGFP gene disrupted by a sense orientation γ-globin intron (intron) and the thymidine kinase polyadenylation signal (pA). The splice donor (SD) and splice acceptor (SA) sites of the intron are indicated. *c*, Types of experimental transgenic lines. The male germ line–specific EGFP retrotransposition cassette was cloned into the 3′ UTR of human L1 elements to create three types of transgenic lines. The transgenes were (i) $L1_{RP}$ under the control of its endogenous promoter (5′ UTR and bent arrow), (ii) $L1_{RP}$ under the control of pPol II and its endogenous promoter and (iii) $L1_{RP}$ containing missense mutations in ORF1 that abolish retrotransposition (approximate location of mutations indicated by asterisks)[4], as a negative control. The number of transgenic lines created for each transgene type is indicated in parentheses.

[1]*Department of Genetics and* [2]*Center for Research on Reproduction and Women's Health, University of Pennsylvania School of Medicine, 475 Clinical Research Bldg., 415 Curie Blvd., Philadelphia, Pennsylvania 19104, USA. Correspondence should be addressed to H.H.K. (e-mail: kazazian@mail.med.upenn.edu).*

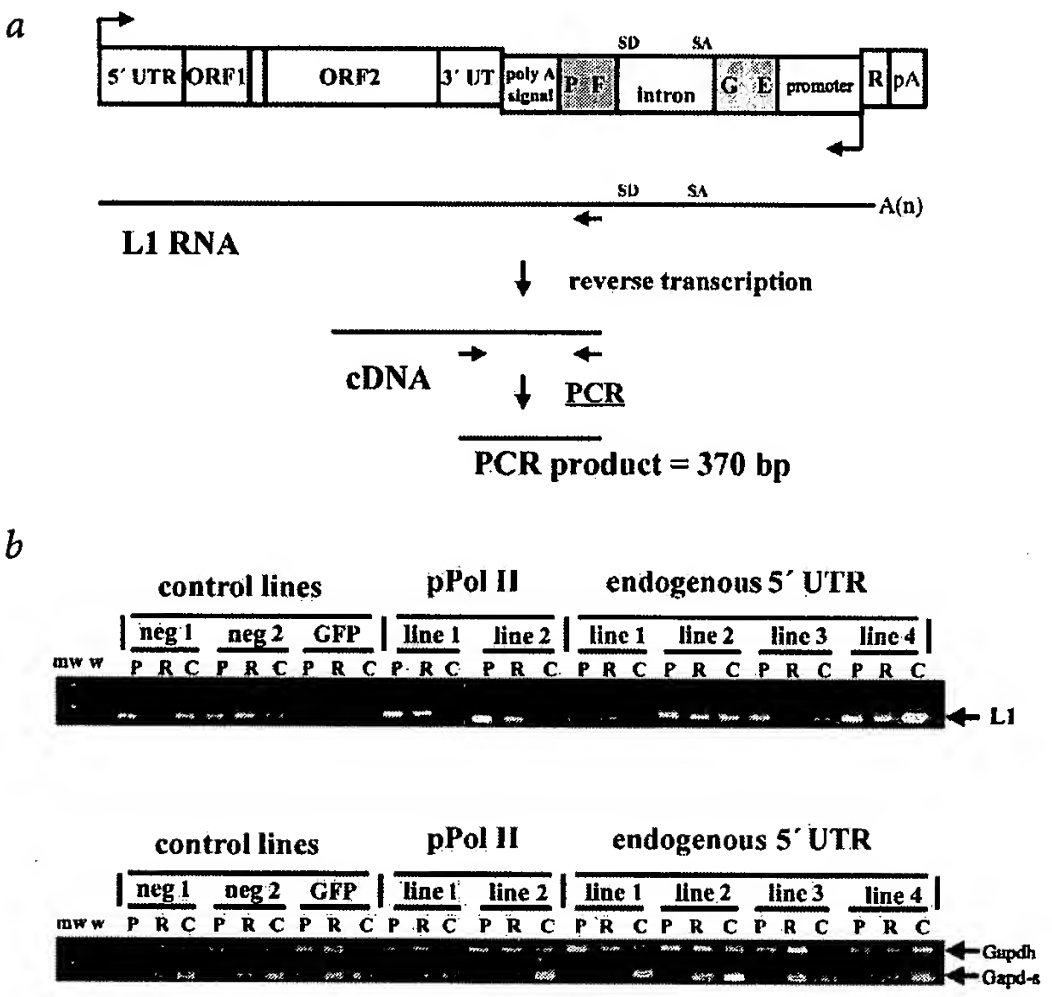

Fig. 2 L1 transgene expression in germ cell fractions. *a*, Strand-specific RT–PCR showing expression of the tagged L1 elements in spermatogenic fractions. We used an oligonucleotide primer specific for the EGFP gene to selectively reverse-transcribe transcripts originating from the L1 promoter. We then added a second oligonucleotide primer specific for the 3′ UTR of the L1 element to carry out PCR. A product of 370 bp is diagnostic of a tagged L1 transcript. *b*, RT–PCR results from the pachytene-spermatocyte (P), round-spermatid (R) and condensing-spermatid (C) fractions from all of our transgenic lines (upper panel). We did RT–PCR using mouse Gapdh primers as a control (lower panel). Lanes containing 1-kb molecular weight marker (mw) and RT–PCR of a water negative control (w) are indicated. P, pachytene spermatocytes; R, round spermatids; C, condensing spermatids.

A cultured-cell assay has been valuable for the study of L1 retrotransposition in a variety of transformed cell lines[4]. But cell culture experiments require retrotransposition from an episome to a chromosome, whereas natural retrotransposition occurs between chromosomal locations. In addition, natural retrotransposition in an organism probably differs from that in transformed cells in the factors used to control the timing and cell-type specificity of transcription as well as in other aspects.

Previous work on mouse L1 elements suggests that L1 expression is germ line–specific. Full-length, sense-strand L1 transcripts have been detected in prepuberal spermatocytes concomitantly with ORF1 protein[5]. Additionally, ORF1 was detected in germ cells of both male and female mice, Leydig cells of embryonic testis and theca cells of adult ovary[5,6] but not in normal somatic tissue. L1 retrotransposition has not, however, been demonstrated experimentally in any cell type in a living organism. Therefore, we created a mouse model of L1 retrotransposition to increase our understanding of L1 biology and to obtain an estimate of the impact of human L1 elements as insertional mutagens.

As a positive control, we created transgenic mice that express enhanced green fluorescent protein (EGFP) in the male mouse germ line and produce sperm with fluorescent acrosomes[7] (Fig. 1*a*). We then cloned an antisense intron into EGFP to create a retrotransposition cassette similar to that used in our cell culture assay[8] (Fig. 1*b*). EGFP transcripts arising from the marker's promoter before retrotransposition contain an antisense intron that cannot be spliced and, therefore, do not produce functional EGFP. A transcript from the L1 promoter that contains the antisense EGFP marker can be spliced, however, thereby removing the intron from the EGFP coding region. Cells express EGFP only after reverse transcription and integration of the intronless cDNA into chromosomal DNA. Therefore, the marker produces functional EGFP only after a retrotransposition event has occurred in male germ cells. We created three types of transgenic lines by cloning the cassette into the 3′ untranslated region (UTR) of three variants of $L1_{RP}$, the most active human element in cell culture[9,10](Fig. 1*c*): (i) $L1_{RP}$, (ii) $L1_{RP}$ with an additional promoter, the mouse RNA polymerase II large subunit promoter (pPol II) and (iii) JM111, an $L1_{RP}$ with two missense mutations in ORF1 that, unlike other L1 mutations characterized to date, completely abolish retrotransposition in cultured cells[4,8].

We purified pachytene spermatocytes, round spermatids and condensing spermatids from each of the mouse lines[11] and carried out strand-specific RT–PCR to verify expression of the L1 transgene (Fig. 2*a*). We detected expression of the tagged L1 transgene in all of the transgenic lines (Fig. 2*b*). We did RT–PCR using primers for mouse glyceraldehyde-3-phosphate dehydrogenase (Gapdh) to ensure that similar amounts of DNA were used in each reaction. Because the Gapdh promoter is silenced as spermatogenesis progresses, the Gapdh band intensity cannot be used to compare RNA amounts from one fraction to

Table 1 • Tissue distribution of L1 transgene expression

Line	Testis	Ovary	Kidney	Lung	Intestine	Liver	Brain
Negative control							
1	***	***	0	0	0	0	0
2	***	**	0	0	0	0	0
Endogenous 5′ UTR							
1	***	***	0	0	0	0	0
2	***	***	0	nd	0	0	0
3	***	**	0	0	0	0	0
4	***	***	0	0	0	0	0
pPol II							
1	***	***	**	**	**	*	*
2	***	nd	0	**	0	0	0

RT–PCR was done on RNA of various tissues of all transgenic lines. 0, no signal; asterisks, signal of varying intensity; nd, results of multiple experiments were equivocal.

another (for example, a pachytene-spermatocyte fraction to a round-spermatid fraction), but it can be used to compare within a fraction (for example, the pachytene-spermatocyte fraction from one line to the pachytene-spermatocyte fraction from another line). Notably, the Gapdh primers also amplify products from Gapd-s, a sperm-specific isoform of Gapdh that increases in abundance as spermatogenesis progresses[12]. Therefore, as Gapdh decreases in intensity, Gapd-s increases, providing an additional confirmation of the germ-cell fraction enrichment. In all RT–PCR experiments, we included RT-minus controls for which we did not detect bands, indicating that the RNA was not contaminated by DNA (data not shown).

In both of the pPol II lines, we found high expression of L1 in early spermatocyte fractions, which diminished in later fractions. This pattern corresponds well to the expression pattern of most mammalian promoters, which are silenced as spermatogenesis progresses[13]. In the lines driven by the L1 promoter alone, however, including the JM111 negative control lines, we found consistent expression throughout spermatogenesis. Notably, the human L1 promoter is expressed highly in male germ cells despite its differences from the various mouse L1 promoters[14–17], suggesting that similar *trans*-acting factors act on both promoters. Expression levels varied between transgenic lines and did not correlate with the estimated transgene copy number (data not shown). Strand-specific RT–PCR showed that transcription from the L1 promoter was restricted to the testis and ovary, whereas the pPol II promoter permitted transcription in other tissues (Table 1).

We first tested for the spliced EGFP mRNA that we expected to occur with retrotransposition by carrying out strand-specific RT–PCR on the condensing spermatid fractions (Fig. 3*a*). As expected, JM111 negative control lines showed no evidence of retrotransposition although they demonstrated expression from the transgene (Fig. 3*b*). In both lines driven by the pPol II promoter and one line (line 4) driven by the endogenous promoter, we detected retrotransposition. Those three lines expressed the highest levels of L1 RNA in pachytene spermatocytes, suggesting that the frequency of retrotransposition correlated with the strength of L1 expression in early germ-cell development. In this assay, functional EGFP transcript could only result from a double-stranded DNA containing the intronless EGFP marker, which could only exist if an EGFP-tagged L1 was transcribed, spliced and reverse-transcribed (that is, if a retrotransposition event had occurred). During L1 retrotransposition, the reverse transcription and chromosomal integration steps occur simultaneously during a coupled reaction termed target-primed reverse transcription[18]. We purified and sequenced several RT–PCR products. The products were of the expected sequence, confirming precise splicing of the intron.

To provide a rough estimate of the frequency of retrotransposition, we diluted RNA from the GFP control line into RNA from a JM111 negative control line to produce a dilution series. We used semi-quantitative RT–PCR to compare the experimental lines to the dilution series, and estimated that retrotransposition occurred in pPol II line 1 at around 1/1,000 spermatids and in other lines at somewhat lower frequencies. These are probably underestimates of retrotransposition frequencies because the amount of GFP RNA produced in a spermatid from a single-copy GFP retrotransposition event is probably lower than that produced from a multi-copy GFP transgene. This predicted reduction in RNA level is supported by the lower fluorescence of positive sperm from the pPol II line compared with positive sperm from the GFP-positive control line (Fig. 3*c*).

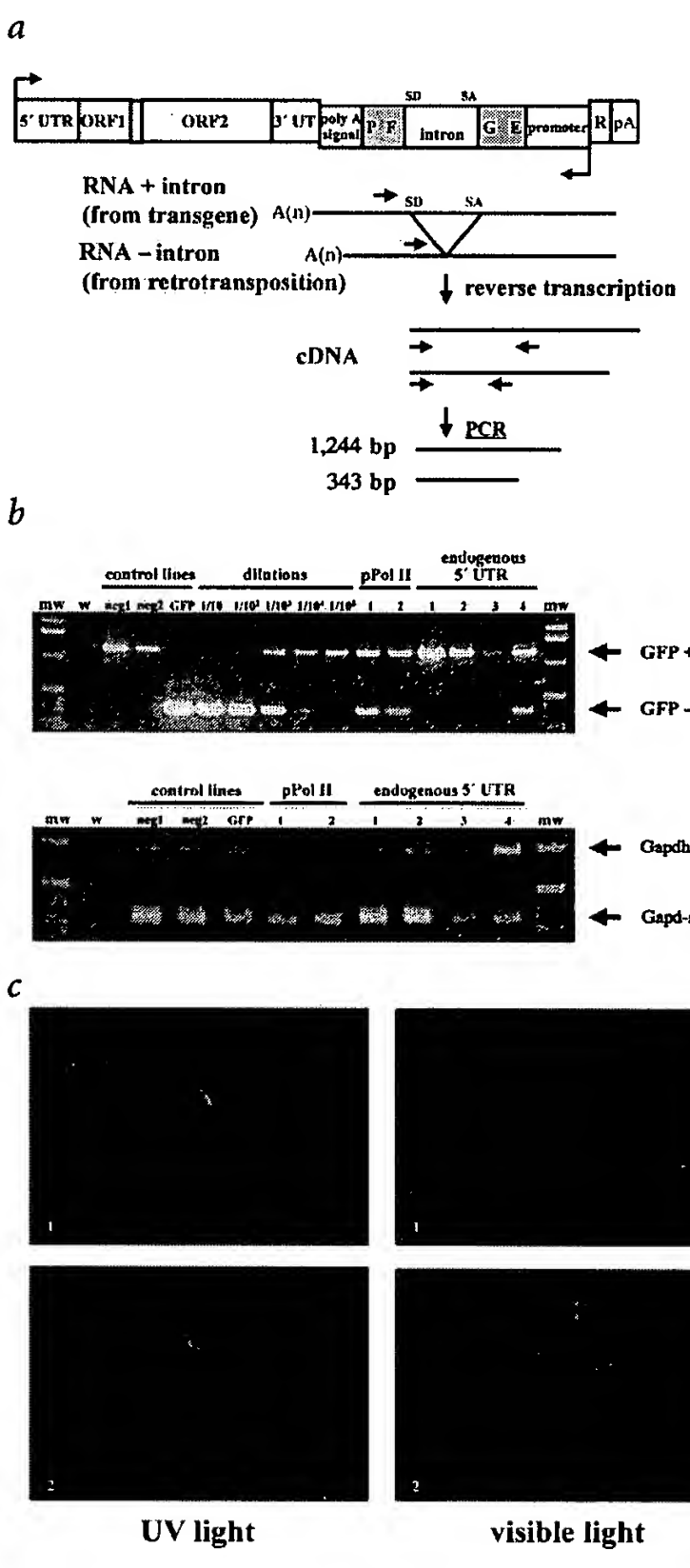

Fig. 3 Expression from retrotransposition events in condensing-spermatid fractions. ***a***, We carried out strand-specific RT–PCR to detect EGFP expression from retrotransposition events. An oligonucleotide primer specific for EGFP was used to selectively reverse-transcribe RNA arising from the preproacrosin promoter. We expected two types of transcripts to arise from this promoter: those with an intron, which come from the transgene (1,244 bp), and those without an intron, which can only arise from a retrotransposition event (343 bp). We then did PCR after addition of an oligonucleotide primer flanking the intron. ***b***, RT–PCR results from the condensing-spermatid fractions from all transgenic lines and the GFP-positive control line (upper panel). To obtain a minimum estimate of the retrotransposition frequency, we diluted condensing spermatid RNA from the GFP control line into condensing spermatid RNA from a negative control line to produce a dilution series. RT–PCR using mouse Gapdh primers was done as a control to ensure that similar amounts of DNA were used in each reaction (lower panel). Lanes containing a 1-kb molecular weight marker (mw) and RT–PCR of a water negative control (w) are indicated. ***c***, Detection of retrotransposition by fluorescence microscopy. We isolated sperm from pPol II line1 and found that some sperm had fluorescent acrosomes when viewed under ultraviolet (UV) light (left panels). On the right are the same sperm under visible light. These pictures were taken with black and white film because it is more sensitive to low light than color film.

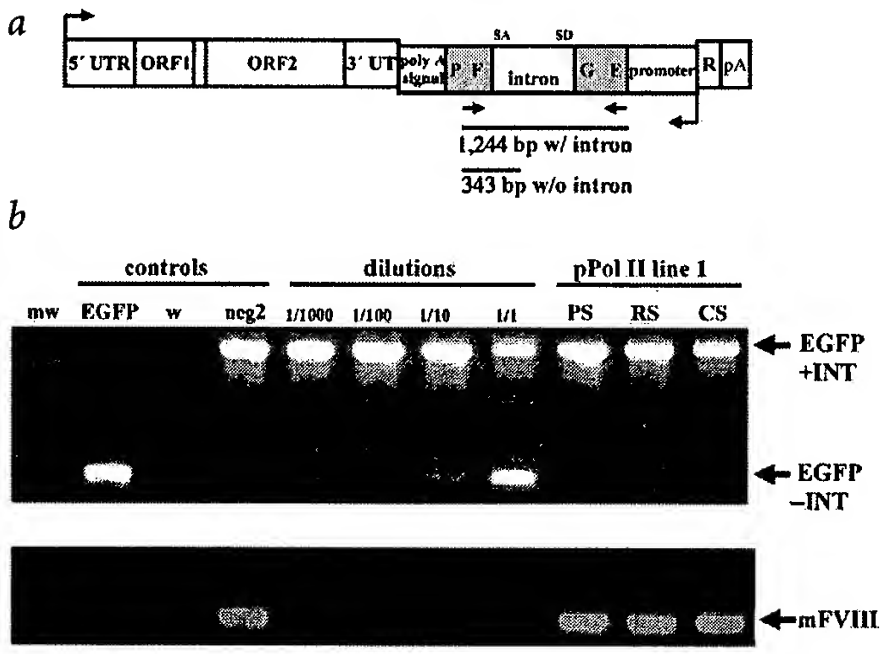

Fig. 4 PCR of germ-cell DNA to detect retrotransposition. *a*, The transgene included EGFP containing a γ-globin intron of 901 bp and would produce a band of 1,244 bp. A retrotransposition event would remove the intron and produce a band of 343 bp. *b*, We did PCR to detect EGFP on genomic DNA isolated from pachytene-spermatocyte (PS), round-spermatid (RS) and condensing-spermatid (CS) fractions from pPol II line 1 (upper panel). We created a dilution series by mixing plasmid DNA containing EGFP with genomic DNA from negative control line 2 (estimated copies of GFP per diploid genome are indicated). We did control PCRs on EGFP-containing plasmid (EGFP), water (w) and genomic DNA from negative control line 2 (neg2). A lane containing a 1-kb molecular weight marker is indicated (mw). We carried out a control PCR on neg2, PS, RS and CS genomic DNAs using mouse factor VIII primers to confirm that similar amounts of DNA were used in each reaction (lower panel).

To improve the estimate of the frequency of retrotransposition, we carried out PCR of genomic DNA from pPol II line 1 spermatogenic fractions (Fig. 4*a*). Using a dilution series consisting of plasmid DNA containing GFP sequences mixed with genomic DNA from a negative control line, we estimated that 1/100 spermatids or more contained a retrotransposition event (Fig. 4*b*).

Realizing that the RT–PCR and PCR results were only order-of-magnitude estimates of the frequency of retrotransposition, we sought formal proof that L1 retrotransposes at a high frequency in this mouse model by characterizing L1 insertions in the offspring of transgenic males of pPol II line 1. We bred F_2 and F_3 pPol II line 1 males and recovered 135 offspring. In these offspring, we found two retrotransposition events. Using inverse PCR, we characterized the insertions and found that they were typical of endogenous L1 retrotransposition events.

One insertion was 1.9 kb in length, contained a poly(A)$^+$ tail of 63 bp, had an inversion and was flanked by target-site duplications of 14 bp (Fig. 5*a*). The inversion occurred at the 5′ end and contained a deletion of 73 bp at the point of inversion. The last three nucleotides at the end of the 5′ target-site duplication were complementary to the nucleotides just proximal to the inversion point on the L1 RNA, suggesting that the inversion was produced by twin priming, a proposed mechanism for the creation of L1 inversions[19]. The predicted cleavage site was typical of L1 endonuclease sites, 5′-TTTT/AA-3′ (refs 20–22). This insertion segregated from the transgene in the subsequent generation.

The second insertion was 4.3 kb in length, contained a poly(A)$^+$ tail of 92 bp and was flanked by target-site duplications of 6 bp (Fig. 5*b*). Again, the predicted cleavage site was typical of that used by the L1 endonuclease. But the mouse that inherited this insertion did not inherit the transgene, not only indicating that retrotransposition almost certainly occurred from one chromosome to another but also strongly suggesting that the event occurred before the end of meiosis I. This result demonstrates that L1 elements can retrotranspose during male gametogenesis.

Our model shows high-frequency chromosome-to-chromosome retrotransposition of a human L1 element in the male germ line of an experimental animal. The *in vivo* insertions were indistinguishable structurally from endogenous L1 insertions in mammals. The observed retrotransposition frequency varied with the genomic context of the transgene and was highly correlated with the amount of L1 transcript present in the pachytene spermatocytes. In three of the four transgenic lines in which the endogenous promoter alone was used to drive L1 transcription, the level of expression in the pachytene spermatocytes was low, and the frequency of retrotransposition was below our limit of detection. In the fourth line, the level of expression in the pachytene spermatocytes was similar to that in the lines driven by the pPol II promoter, and the frequency of retrotransposition was roughly one-half that of pPol II line 1, as estimated by RT–PCR (Fig. 3*b*). We also found a correlation between L1 expression level and retrotransposition frequency in experiments using cultured cells (data not shown).

Among the 40–80 potentially active L1 elements in the human genome, some are older and common to all humans, whereas others are younger, and their presence is polymorphic. The younger elements are more active in cell culture[2,10,23], though most are far less active than $L1_{RP}$ owing to mutations that attenuate their ability to retrotranspose. The few very active elements are highly polymorphic. Therefore, each individual has a different complement of such L1 elements, only some of which reside in genomic regions that permit expression. This raises the question of how such a small number of potentially active L1 elements can impose such a high mutational load on the genome. Prior crude estimates of the frequency of L1 retrotransposition

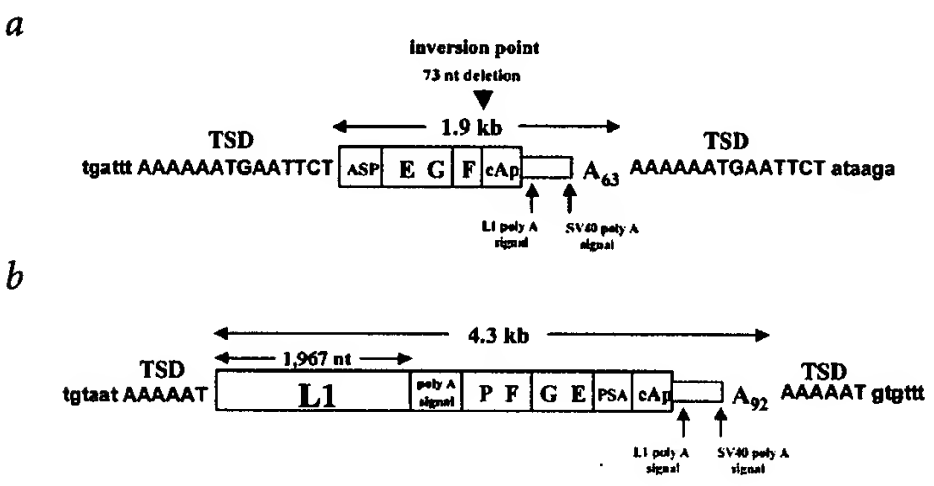

Fig. 5 L1 insertions in transgenic mice. *a*, Insertion #1 was 1.9 kb in length and contained an inversion of the 5′ end with a deletion of 73 bp at the inversion point, a common occurrence in L1 retrotransposition. A 63-bp poly(A)$^+$ tail was added after the SV40 poly(A)$^+$ signal, and the insertion was flanked by 14-bp target-site duplications (TSDs, sequence shown in capital letters). Flanking sequence is shown in lower-case letters. *b*, Insertion #2 was 4.3 kb in length and contained a 92-bp poly(A)$^+$ tail added after the SV40 poly(A)$^+$ signal, and the insertion was flanked by 6-bp target-site duplications (TSDs, sequence shown in capital letters). The L1 inserted into intron 1 of a predicted gene (mCG57584 from Celera Discovery System) on chromosome 9. Comparison of insertions #1 and #2 with their respective empty site sequences showed neither added nor deleted nucleotides other than the target-site duplications.

based on mutation rates in specific genes and overall mutation rates in germ cells ranged from 1 in 12 to 1 in 140 haploid genomes[24,25]. Our data suggest that a single very active human L1 can retrotranspose at a rate that is well within the range of previous empirical estimates when residing in a genomic locus that permits a high rate of transcription during spermatogenesis.

Our L1 retrotransposon mouse model will be valuable in answering a number of questions concerning the biology of mammalian retrotransposition. Is L1 retrotransposition in mice suppressed after multiple generations, analogous to retrotransposition of I factor in *Drosophila melanogaster*[26]? What are the factors that allow transcription and retrotransposition of L1 elements in the germ line, but suppress somatic cell transcription? Do defects in methylation or DNA double-strand–break repair increase the frequency of retrotransposition in germ cells? Human L1 transgenes may also be useful as random insertional mutagens for determining gene function in mice.

Methods

Cloning of transgenes. We created clone pBSKS-AcEGFP-INT (a male germ line–specific EGFP retrotransposition cassette) by carrying out a three-way ligation with an approximately 2,960-bp fragment from pBluescript KS(–) (Stratagene), a 453-bp *XhoI*/*SalI* fragment from pGES36-TKpolyA (a plasmid containing the EGFP retrotransposition cassette with a thymidine kinase poly(A)$^+$ signal) and a 2,395-bp *Hin*dIII/*Xho*I fragment from pRJD538. We cloned the AcEGFP cassette as an *Xma*I/blunted *Sal*I fragment into the 3′ UTR of $L1_{RP}$ (pJCC5-$L1_{RP}$ (ref. 8) cut with *Xma*I/*Bst*Z17I) to create pBS-$L1_{RP}$-AcEGFP or into the 3′ UTR of $L1_{RP}$(JM111), the $L1_{RP}$ element containing two missense mutations in the ORF1 coding region, (pJCC5-$L1_{RP}$(JM111)) cut with *Xma*I/*Bst*Z17I) to create pBS-$L1_{RP}$(JM111)-AcEGFP. We cloned the tagged $L1_{RP}$ from pBS-$L1_{RP}$-AcEGFP as a *Not*I/blunted *Apa*I fragment into the multiple cloning site of pRJD099 (promoter-less CEP4-based vector) cut with *Not*I/blunted *Sfi*I to create p99-$L1_{RP}$-AcEGFP or into the *Not*I/blunted *Apa*I–digested multiple cloning site of pRJD907 (CEP4-based vector containing the mouse pPol II promoter in place of the CMV promoter) to create p907-$L1_{RP}$-AcEGFP. We similarly cloned the tagged $L1_{RP}$(JM111) element from pBS-$L1_{RP}$(JM111)-AcEGFP into pRJD099 to create p99-$L1_{RP}$(JM111)-AcEGFP. We purified transgenes using Elutip-D mini-column (Schleicher & Schuell). The purified transgenes were microinjected into fertilized mouse oocytes (strain B6SJF1) by the University of Pennsylvania School of Medicine Transgenic & Chimeric Mouse Core Facility.

L1 expression assay. To detect expression of the tagged L1 transgene, we carried out a strand-specific RT–PCR with the OneStep kit (Qiagen) in 50 μl volume using 500 ng of RNA per reaction. We added oligonucleotide 1239+ during the reverse transcriptase step (30 min at 50 °C) to selectively reverse-transcribe transcripts arising from the L1 promoter (as opposed to the EGFP promoter) and then added oligonucleotide L16045(30) before the PCR step. Our PCR conditions included an initial step at 95 °C (15 min), 30 cycles of amplification (30 s at 94 °C, 45 s at 50 °C, 1 min at 72 °C) and a final step at 72 °C (10 min). We did RT-minus controls by setting up identical reactions in parallel and leaving the reactions on ice during the reverse transcription step. We used a PTC-200 Peltier Thermal Cycler (MJ Research) for RT–PCR. We digested samples for 30 min at 37 °C with 0.5 μl (250 ng) DNase-free RNase (Roche) and then analyzed samples by separating aliquots on a 1.0% agarose gel. We did control RT–PCRs using mouse Gapdh-specific oligonucleotides mGAPDH5′ and mGAPDH3′. We did 30 cycles of amplification (30 s at 94 °C, 45 s at 58 °C, 1 min at 72 °C) followed by a final step at 72 °C for 10 min. All primer sequences are available upon request.

Detection of retrotransposition events by RT–PCR. To detect expression from the EGFP gene of retrotransposition events, we carried out strand-specific RT–PCR as described above, except that we added oligonucleotide GFP1310R during the reverse transcriptase step to selectively reverse-transcribe transcripts arising from the EGFP promoter. We added oligonucleotide GFP968F before the PCR step. Our PCR conditions included an initial step at 94 °C (10 min), 35 cycles of amplification (10 s at 94 °C, 30 s at 66.8 °C, 30 s at 72 °C) and a final step at 72 °C (10 min). We did RT-minus controls by setting up identical reactions in parallel and leaving the reactions on ice during the reverse transcription step.

DNA sequence analysis. We used 35 ng of RT–PCR product and 3.2 pmol oligonucleotide primer GFP968F for sequencing reactions, which were done using ABI 377 and 373A Stretch sequencers (DNA sequencing core, University of Pennsylvania).

PCR of spermatogenic DNA. We isolated genomic DNA from spermatogenic germ-cell fractions with a Blood and Cell Culture DNA Mini Kit (Qiagen). To make a standard reference, we diluted pEGFP-N1 into round-spermatid fraction genomic DNA from negative control line 2 at 1/1,000, 1/100, 1/10 and 1 copies per haploid genome. We carried out PCR on 500 ng genomic DNA from all fractions with EGFP-specific primers GFP968F and GFP1310R. We did amplifications in 50 μl containing 1.25 U *Taq* DNA polymerase (Roche), 1× PCR reaction buffer (Roche), 0.2 mM of each dNTP, 200 ng of each oligonucleotide primer and approximately 500 ng genomic DNA. Our PCR conditions included an initial step at 94 °C (10 min), 35 cycles of amplification (10 s at 94 °C, 30 s at 66.8 °C, 30 s at 72 °C) and a final step at 72 °C (10 min). We did control PCR on genomic DNA from pachytene-spermatocyte, round-spermatid and condensing-spermatid fractions from pPol II mouse line 1 and control round-spermatid fraction genomic DNA from negative control mouse line 2 using mouse factor VIII-specific primers. The oligonucleotides used for PCR were MC-18 and MC-19. We did amplifications in 50 μl containing 1.25 U *Taq* DNA polymerase (Roche), 1× PCR reaction buffer (Roche), 0.2 mM of each dNTP, 200 ng of each oligonucleotide primer and approximately 500 ng genomic DNA. Our PCR conditions included an initial step at 94 °C (10 min), 35 cycles of amplification (10 s at 94 °C, 30 s at 58.4 °C, 1 min at 72 °C) and a final step at 72 °C (10 min). We analyzed samples by separating aliquots on a 1.0% agarose gel.

Transgenic mice. We bred all founder mice with non-transgenic 129/SV mice (Jackson Laboratories) to establish stable transgenic lines. In subsequent matings, we bred transgenic mice with non-transgenic 129/SV mice. The transgenic mouse protocol, titled "A human transposable element", was reviewed by the Institutional Biosafety Committee of the University of Pennsylvania and was approved on 28 April 2000. The protocol was reviewed by the Institutional Animal Care and Use Committee of the University of Pennsylvania and approved on 24 July 2000.

Characterization of insertions by inverse PCR. We digested genomic DNA with either *Sph*I (for insertion #1) or *Afl*II (for insertion #2) and carried out self-ligations on restricted fragments at low DNA concentrations. After ethanol precipitation, we subjected 400 ng of ligated DNA to an initial round of PCR using the primer set 1EGFPBRDB1 and 2EGFPDOWN with the Expand Long Template PCR System (Roche). We used a 2-μl aliquot from this reaction in a second-round nested PCR reaction using the primer sets 3EGFPUP and 5NEOGFPDOWNB for insertion #1 and 3EGFPUP and 4EGFPDWN1 for insertion #2. We isolated PCR products following gel electrophoresis. We obtained the sequence flanking the 3′ end of the L1 inserts using an oligonucleotide annealing at the end of the L1 poly(A)$^+$ tail, 5′ $A_{23}T$ 3′. We identified insertion sites for insertion #1 and #2 in entries GA_x5J8B7TT2F0 and GA_x5J8B7W3MYM, respectively, of the Celera Discovery System mouse genome database. Based on this sequence, we amplified insertion #2, together with about 320 nt of 5′ flanking DNA, using primers situated in the EGFP cassette (3EGFPUP) and flanking region (4AFL1AT5P2). We also amplified the entire insertion #1 using primers in the 5′ and 3′ flanking DNA, B141SPH3P2 and D141SPH5P4. We cloned PCR fragments into the vector pCR2.1 (Invitrogen) and sequenced them in their entirety.

Acknowledgments

We thank A. Farley, K.-S. Kim and M.C. Cha for technical assistance; the Transgenic and DNA Sequencing Cores of the University of Pennsylvania for generation of transgenic mouse lines and for DNA sequences, respectively; and J. Moran and B. Brouha for critical reading of the manuscript. E.M.O. was supported by a Howard Hughes Predoctoral Fellowship, and H.H.K. was supported by grants from the US National Institutes of Health.

Competing interests statement

The authors declare that they have no competing financial interests.

Received 18 April; accepted 24 September 2002.

1. The International Human Genome Sequencing Consortium. *Nature* **409**, 860–921 (2001).
2. Ostertag, E.M. & Kazazian, H.H. Jr. Biology of mammalian L1 retrotransposons. *Ann. Rev. Genet.* **35**, 501–538 (2001).
3. Sassaman, D.M. *et al.* Many human L1 elements are capable of retrotransposition. *Nature Genet.* **16**, 37–43 (1997).
4. Moran, J.V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–927 (1996).
5. Branciforte, D. & Martin, S.L. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell. Biol.* **14**, 2584–2592 (1994).
6. Trelogan, S.A. & Martin, S.L. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Natl Acad. Sci. USA* **92**, 1520–1524 (1995).
7. Nakanishi, T. *et al.* Real-time observation of acrosomal dispersal from mouse sperm using GFP as a marker protein. *FEBS Lett.* **449**, 277–283 (1999).
8. Ostertag, E.M. *et al.* Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.* **28**, 1418–1423 (2000).
9. Schwahn, U. *et al.* Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.* **19**, 327–332 (1998).
10. Kimberland, M.L. *et al.* Full-length human L1 insertions retain the capacity for high-frequency retrotransposition in cultured cells. *Hum. Mol. Genet.* **8**, 1557–1560 (1999).
11. Romrell, L.J., Bellvé, A.R. & Fawcett, D.W. Separation of mouse spermatogenic cells by sedimentation velocity. A morphological characterization. *Dev. Biol.* **49**, 119–131 (1976).
12. Welch, J.E., Schatte, E.C., O'Brien, D.A. & Eddy, E.M. Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mouse spermatogenic cells. *Biol. Reprod.* **46**, 869–878 (1992).
13. Steger, K. Transcriptional and translational regulation of gene expression in haploid spermatids. *Anat. Embryol.* **199**, 471–487 (1999).
14. Swergold, G.D. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* **10**, 6718–6729 (1990).
15. Loeb, D.D. *et al.* The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol. Cell. Biol.* **6**, 168–182 (1986).
16. Padgett, R.W., Hutchison, C.A. 3rd & Edgell, M.H. The F-type 5' motif of mouse L1 elements: a major class of L1 termini similar to the A-type in organization but unrelated in sequence. *Nucleic Acids Res.* **16**, 739–749 (1988).
17. Naas, T.P. *et al.* An actively retrotranposing, novel subfamily of mouse L1 elements. *EMBO J.* **17**, 590–597 (1998).
18. Luan, D.D., Korman, M.H., Jakubczak, J.L. & Eickbush, T.H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposable elements. *Cell* **72**, 595–605 (1993).
19. Ostertag, E.M. & Kazazian, H.H. Jr. Twin Priming, a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genet. Res.* **11**, 2059–2065 (2001).
20. Feng, Q., Moran, J.V., Kazazian, H.H. Jr. & Boeke, J.D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).
21. Jurka, J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl Acad. Sci. USA* **94**, 1872–1877 (1997).
22. Cost, G.J. & Boeke, J.D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochem.* **37**, 18081–18093 (1998).
23. Boissinot, S., Chevret, P. & Furano, A.V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol. Biol. Evol.* **17**, 915–928 (2000).
24. Kazazian, H.H. Jr. An estimated frequency of endogenous insertional mutations in humans. *Nature Genet.* **22**, 130 (1999).
25. Li, X. *et al.* Frequency of recent retrotransposition events in the human factor IX gene. *Hum. Mutat.* **17**, 511–519 (2001).
26. Jensen, S., Gassama, M.P. & Heidmann, T. Taming of transposable elements by homology-dependent gene silencing. *Nature Genet.* **21**, 209–212 (1999).
27. Hohjoh, H. & Singer, M.F. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J.* **15**, 630–639 (1996).
28. Kolosha, V.O. & Martin, S.L. *In vitro* properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc. Natl Acad. Sci. USA* **94**,10155–10160 (1997).
29. Mathias, S.L., Scott, A.F., Kazazian, H.H., Boeke, J.D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* **254**, 1808–1810 (1991).
30. Fanning, T. & Singer, M. The line-1 DNA-sequences in 4 mammalian orders predicts proteins that conserve homologies to retrovirus proteins. *Nucleic Acids Res.* **15**, 2251–2260 (1987).